

IN THE UNITED STATES  
PATENT AND TRADEMARK OFFICE

JCS 11/13/98  
PATENT APPLICATION

INVENTORS: Jiashu CHEN; and Christopher Anton WENDT  
CASE: CHEN 3-1  
FILE: METHOD AND APPARATUS FOR PROCESSING  
INTERAURAL TIME DELAY IN 3D DIGITAL AUDIO

JCS 11/13/98  
09/190208  
PTO

11/13/98  
Barcode

ASSISTANT COMMISSIONER FOR PATENTS  
WASHINGTON, D.C. 20231

SIR:

Enclosed are the following papers relating to the above-named application for patent:

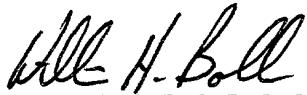
Specification (including claims and Abstract) - 17 pages  
6 Informal sheets of drawing(s)

CLAIMS AS FILE				
	NO. FILED	NO. EXTRA	RATE	CALCULATIONS
Total Claims	14 - 20 =	0	x \$22 =	\$0
Independent Claims	3 - 3 =	0	x \$82 =	\$0
Multiple Dependent Claim(s), if applicable			\$270 =	\$0
Basic Fee				\$790
TOTAL FEE:				\$790

Please file the application and charge **Lucent Technologies Deposit Account No. 12-2325** the amount of **\$790**, to cover the filing fee. Duplicate copies of this letter are enclosed. In the event of non-payment or improper payment of a required fee, the Commissioner is authorized to charge or to credit **Deposit Account No. 12-2325** as required to correct the error.

Please address all correspondence to **FARKAS & MANELLI, PLLC, 2000 M Street, N.W. Suite 700, Washington, DC 20036-3307**, and all telephone calls to William H. Bollman at his Washington, DC local number of (202) 261-1000.

Respectfully submitted,



William H. Bollman  
Reg. No.: 36,457  
Attorney for Applicant(s)

Date: November 13, 1998

**Farkas & Manelli, PLLC**  
**2000 M Street, N.W. Suite 700**  
**Washington, DC 20036-3307**  
**(202) 261-1000**

# APPLICATION UNDER UNITED STATES PATENT LAWS

Invention: **METHOD AND APPARATUS FOR PROCESSING  
INTERAURAL TIME DELAY IN 3D DIGITAL AUDIO**

Inventor(s): Jiashu CHEN; and  
Christopher Anton WENDT

Farkas & Manelli P.L.L.C.  
2000 M Street, N.W.  
7<sup>th</sup> Floor  
Washington, D.C. 20036-3307  
Attorneys  
Telephone: (202) 261-1000

This is a:

- Provisional Application
- Regular Utility Application
- Continuing Application
- PCT National Phase Application
- Design Application
- Reissue Application
- Plant Application

## SPECIFICATION

# METHOD AND APPARATUS FOR PROCESSING INTERAURAL TIME DELAY IN 3D DIGITAL AUDIO

This application claims priority from U.S. Patent Application  
5 No. 60/065,855 entitled "Multipurpose Digital Signal Processing System"  
filed November 14, 1997, the specification of which is explicitly  
incorporated herein by reference.

## BACKGROUND OF THE INVENTION

### 10 1. Field of the Invention

This invention relates generally to three dimensional (3D) sound. More particularly, it relates to a digital implementation of interaural time delays used in 3D digital sound applications.

### 15 2. Background of Related Art

Many high-end consumer devices provide the option for three-dimensional (3D) sound, allowing a more realistic experience when listening to sound. In some applications, 3D sound allows a listener to perceive motion of an object from the sound played back on a 3D audio  
20 system.

Atal and Schroeder established cross-talk canceler technology as early as 1962, as described in U.S. Patent No. 3,236,949, which is explicitly incorporated herein by reference. The Atal-Schroeder 3D sound cross-talk canceler was an analog implementation using  
25 specialized analog amplifiers and analog filters. To gain better sound positioning performance using two loudspeakers, Atal and Schroeder included empirically determined frequency dependent filters. Without doubt, these sophisticated analog devices are not applicable for use with today's digital audio technology.

Interaural time difference (ITD), i.e., the difference in time that it takes for a sound wave to reach both ears, is an important and dominant parameter used in 3D sound design. The interaural time difference is responsible for introducing binaural disparities in 3D audio or 5 acoustical displays. In particular, when a sound object moves in a horizontal plane, a continuous interaural time delay occurs between the instant that the sound object impinges upon one of the ears and the instant that the same sound object impinges upon the other ear. This ITD is used to create aural images of sound moving in any desired direction 10 with respect to the listener.

The ears of a listener can be 'tricked' into believing sound is emanating from a phantom location with respect to the listener by appropriately delaying the sound wave with respect to at least one ear. This typically requires appropriate cancellation of the original sound wave 15 with respect to the other ear, and appropriate cancellation of the synthesized sound wave to the first ear.

Atal-Schroeder implemented the delays and cancellations with appropriate analog filters and analog amplifiers, as shown herein in Figs. 5 and 6. Figs. 5 and 6 herein are described in detail in the Atal- 20 Schroeder U.S. Pat. No. 3,236,949 with reference therein to Figs. 2 and 4, respectively. Fig. 5 herein shows the conventional 3D sound system for creating the image of sound from a phantom locality with respect to the listener, while Fig. 6 herein shows the analog delay line with multiple tap points implemented by Atal-Schroeder.

25 Thus, the interaural time delay is manipulated to synthesize localities of the source of particular sounds, and to create the sense of motion of particular sounds.

Conventional 3D sound systems embed the interaural time difference in empirically determined head-related transfer functions 30 (HRTFs), typically determined with a mannequin head implanted with

microphones in its ears. The available delays typically have a relatively large resolution, e.g., 100 microseconds, formed by null filter taps, as disclosed by Atal-Schroeder.

However, there are at least two basic problems with the  
5 implementation of the conventional analog approach in a digital environment. First of all, the large resolution in the available time delays cause discretely sampled interaural time differences for the expected position of a listener. Thus, a 'closest' or 'best fit' ITD must be chosen, which may be up to 50% away from the ideal parameter. This may cause  
10 a jittering effect in the sense of movement of the sound by the listener. Moreover, implementation of a digital filter emulating the analog filter having multiple taps as shown herein in Fig. 6 is computationally involved, providing a level of system inefficiency from a computational view.

One conventionally proposed implementation of a digital 3D  
15 sound system to provide a more accurate ITD based on the given resolution has been to interpolate the entire HRTF set such that the ITD becomes interpolated as well. Unfortunately, interpolation itself can become a computationally intense requirement which likely adds to, rather than cures, the computational inefficiency otherwise associated with  
20 digital 3D sound systems.

There is thus a need for an efficient and simplified method and apparatus for providing digital 3D sound.

## **SUMMARY OF THE INVENTION**

25 In accordance with the principles of the present invention, a digital delay line for use in a 3D audio sound system comprises a first delay module providing a choice of any delay within a first resolution. A second delay module is in series with the first delay module. The second delay module provides a choice of any of a plurality of additional fractional

delays. Each of the additional fractional delays is less than the first resolution.

A method for providing an interaural time delay in a digital 3D sound system in accordance with another aspect of the present invention comprises selecting one of a plurality of available first time delays having a first resolution between each of the plurality of available first time delays. Additionally, one of a plurality of available second time delays is selected. Each of the plurality of available second time delays is less than the first resolution. The selected first time delay is added to the second time delay to provide a desired interaural time delay.

#### **BRIEF DESCRIPTION OF THE DRAWINGS**

Features and advantages of the present invention will become apparent to those skilled in the art from the following description 15 with reference to the drawings, in which:

Fig. 1 is a block diagram showing the digital 3D sound system including a digital interaural delay line, in accordance with the principles of the present invention.

Fig. 2 is a more detailed diagram showing the digital 3D 20 sound system for creating 3D sound in a digital environment, in accordance with the principles of the present invention.

Fig. 3 is a diagram showing the implementation of multiple digital audio streams using a common bank of fractional delay filters, in accordance with the principles of the present invention.

Fig. 4 shows a process for creating an improved ITD look-up 25 table suitable for use in an ITD look up table for use with 3D sound applications as shown in Figs. 1 and 2, in accordance with the principles of the present invention.

Fig. 5 shows a conventional 3D sound system for creating 30 the image of sound from a phantom locality with respect to the listener.

Fig. 6 shows a conventional analog delay line with multiple tap points implemented by Atal-Schroeder.

## DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

5 In accordance with the principles of the present invention, the ITD is extracted from measured and empirically determined HRTFs, smoothed, and implemented in a look-up table. Implementation of the ITD is provided by a delay line including both an integer portion providing rough estimate delays and a fractional portion providing a very accurate  
10 10 delay and eliminating discontinuities in the listening field to provide a more relaxed listening sweet spot.

The present invention provides a digital filter bank having a simple and inherently low cost architecture for performing a stable cross-talk cancellation, providing excellent localization and externalization of  
15 15 virtual sound images.

In accordance with the principles of the present invention, head-related transfer functions corresponding to the speaker positions are recorded and used to construct the filter coefficient. The relationship between the speaker position and filter design were studied to provide a  
20 20 more relaxed listening “sweet spot” where the 3D sound effects are optimized. Thus, the listener does not have to sit in a very accurately placed position with respect to the loudspeakers to appreciate the 3D aspects of the audio rendered by only two loudspeakers.

Fig. 1 is a block diagram showing the basic components of  
25 25 the disclosed embodiment of a digital 3D sound system including a digital interaural time delay line, in accordance with the principles of the present invention.

In particular, a sound source **220** is input into a digital interaural time delay line **254**. the interaural delay line **254** includes an  
30 30 integer delay module **250** providing a rough estimate of the desired

interaural time delay, and a fractional delay module **252** providing a highly refined additional time delay. In the disclosed embodiment, both the particular settings of both the integer delay module **250** and the fractional delay module **252** are chosen from among a plurality of predetermined 5 delays, greatly reducing or eliminating the otherwise intensive calculations necessary to interpolate a particular interaural time delay.

The particular delay associated with the left (or right) ear signal **260** and the right (or left) ear signal **262** providing the desired localization of the sound image is provided by a localization control 10 module **270**.

Fig. 2 is a more detailed diagram showing the digital 3D sound system shown in Fig. 1.

In particular, the integer delay module **250** of the disclosed embodiment is comprised of a first-in, first-out (FIFO) buffer **204**. The 15 FIFO buffer **204** may be of any suitable width, e.g., 16 bits, corresponding to the length of the digital audio samples. Moreover, the length of the FIFO buffer **204** will be based on the largest delay necessary to implement the desired 3D sound imaging. The particular delay is related to the selected number of clock cycles after the particular digital audio 20 sample was input to the FIFO buffer **204**. This selection of an integer delay time is represented in Fig. 2 with a multiplex switch **206**. The use of any of the particular digital audio samples **224a-224d** are fed serially into the FIFO buffer **204**, with the arrows from each of the samples **224a-224d** representing tap numbers.

25 The clock cycle of the FIFO buffer **204** relates to one over the sample rate. Thus, with an exemplary sample rate of 22 kiloHertz, the 'integer' portion, or resolution of the integer delay module **250** is 1/22,000 or approximately 45 microseconds (uS).

The second portion of the digital interaural delay line **254** 30 provides a much more refined 'fractional' delay with a fractional delay

module **252**. This fractional delay is provided by the selection of any one of a plurality of fractional delay filters **208-212**.

The fractional delay module **252** effectively produces an adjustable digital delay with a finer resolution than the integer delay module **250**. Each of the fractional delay filters **208-212** is a so-called all-pass filter that has a variable phase shift, corresponding to the required fractional delay. The number of phases (i.e., fractional delay filters **208-212**) is determined empirically by behavioral testing of human listening.

In the disclosed embodiment, 64 fractional delay filters are utilized, each providing an incrementally greater delay, in finely resolved increments suitable to the application. For instance, at the exemplary sample rate of 22 kiloHertz, the resolution between the fractional delay filters **208-212** is (45 uS)/64, or about 0.7 uS resolution. This particular fine resolution (and the rough estimate resolution provided by the integer delay module **250**) can be adjusted based on the needs of the particular application.

Each fractional delay filter **208-212** is a finite impulse response (FIR) filter, i.e., a polyphase filter, effecting the desired delay. Each of the fractional delay filters **208-212**, and/or the fractional delay controlled switch **216** and/or the multiplexer **214** can be implemented in any suitable processor, e.g., in a digital signal processor (DSP), microprocessor, or microcontroller. Alternatively, the digital filters can be implemented in hardware in accordance with the principles of the present invention.

In the exemplary embodiment utilizing a sampling rate of 22 kiloHertz, the first fractional delay filter **208** provides 0.7 uS delay to a digital audio sample passing therethrough, the second fractional delay filter **210** provides approximately 1.4 uS delay, etc., until the last fractional delay filter **212** which provides approximately 44.3 uS delay.

Selection of the appropriate fractional delay filter **208-212** is implemented by a multiplexer **214** in the fractional delay module **252**. In the shown embodiment, the fractional delay filters **208-212** are each implemented in a processor, e.g., in a digital signal processor, and 5 selection of an appropriate one of the fractional delay filters **208-212** is desirable at the front end to avoid wasted computational power by running fractional delay filters **208-212** which are not being used for that particular audio sample.

The interaural time delay is controlled by the localization 10 control module **270**, which includes a 3D audio application source position controller **222**, an interaural time delay (ITD) look-up table **220**, and an integral and fractional delay selector **218**. In the disclosed embodiment, the localization control module **270** is implemented in a suitable processor, e.g., in a microprocessor, microcontroller, or digital signal 15 processor (DSP). Of course, the localization control module **270** may alternatively be partially or wholly implemented in hardware, e.g., using programmable array logic.

The 3D audio application source position control **222** selects a desired 'phantom' position of the sound sample currently being input to 20 the digital interaural delay line **254**. The desired location may have a desired x, y and z coordinate with respect to a reference point, e.g., the center of the listener's head. Based on the desired location, an associated ITD is determined in the ITD look-up table **220**. The integral and fractional delay selector determines the largest integer value which 25 can be achieved within the resolution of the integer delay module **250** without exceeding the desired ITD, and appropriately controls the integer delay module **250** to provide that desired delay to the audio sample. Similarly, the remainder or fractional portion of the desired ITD which is not provided by the integer delay module **250** is provided by an

appropriate selection of a desired one of the available fractional delay filters **208-212** in the fractional delay module **252**.

Fig. 3 is a diagram showing the implementation of multiple digital audio streams using a common bank of fractional delay filters, in accordance with the principles of the present invention. Thus, the plurality of fractional delay filters **208-212** can be utilized by a plurality of audio sources for the same listener, avoiding the need to duplicate the fractional delay module **252** for each audio source.

Fig. 4 shows a process for creating the ITD look-up table **220** shown in Fig. 2.

In particular, in step **102**, binaural impulse responses are empirically measured with a sound source at various locations around the listening environment, e.g., at incremental points along a sphere about the sound source.

In step **104**, the ITD information is extracted from the empirically measured information obtained in step **102**, and a ‘mesh’ of ITD values for each appropriate point on the sphere is determined. In particular, the ITD samples may be extracted from measured left-right ear head-related transfer functions (HRTFs) using cross-correlation. These samples can be viewed as discrete samples of an underline continuous ITD function of azimuth and elevation coordinates.

In step **106**, to avoid the ‘jittering’ and other undesirable effects for the listener, the ITD mesh determined in step **104** is smoothed using any appropriate smoothing algorithm. For instance, the ITD samples may be regularized using a “generalized spline model” or appropriately filtered and interpolated by a two-dimensional filter to gain smoothness and continuity. While this smoothing may be calculation intensive, it is performed once, off-line, and not performed in real-time as digital audio samples are received.

In step 108, the smoothed ITD mesh is input into the ITD look-up table 220. The ITD mesh may utilize any appropriate coordinate system, e.g., spherical coordinates or a standard x, y and z coordinate system.

5 In the disclosed embodiment it was determined that the finest time resolution of the overall delay, i.e., the combination of the delay provided by the integer delay module 250 and the fractional delay module 252, is preferably less than 1 microsecond ( $\mu$ S) such that any discontinuity caused in the sound stream is under the hearing threshold of  
10 a typical human. In the case of a high sampling rate, faster time resolution may be preferred. For example, with a 22.05 kiloHertz sampling rate of an audio stream, a 64-phase polyphase filter was used to obtain sub-microsecond resolution in the time delay. In another example, a 60-phase polyphase filter was used to provide the necessary time  
15 delays for a suitable presentation of a audio stream sampled at 48 kiloHertz.

While the fractional delay filters 208-212 in the disclosed embodiment are each a FIR (polyphase) filter, the principles of the present invention are equally applicable to the use of other filters or digital  
20 delays which provide the required delay in a digital audio sample.

The digital interaural delay line 254 in accordance with the principles of the present invention can be implemented in any suitable processor or computer system. For instance, the digital interaural delay line 254 can be implemented at a host level in a personal computer (PC)  
25 based platform using regular instruction sets or MMX™ technology, or can be implemented in a digital signal processor (DSP).

To further improve upon efficiency in accordance with the principles of the present invention, the delay may be fixed for one ear, and varied for the sound intended for the other ear, according to the desired  
30 movement of the source sound. This alternative method may save as

many as half of the instruction cycles required to otherwise process a variably delayed sound to both ears.

The appropriately delayed left and right ear signals can be forwarded to a next stage for further processing, or sent directly to 5 headphones or loudspeakers for presentation to the listener, as a simple binaural signal processing method.

Thus, in accordance with the principles of the present invention, a solution to the problem of generating a proper interaural time delay in 3D audio and acoustical virtual display applications is 10 implemented with the requirement for little processing delay. The principles of the present invention saves instruction cycles of a processor over conventional interpolation techniques, and use of the FIFO buffer 204 eliminates the need for the storage of a suitable plurality of null taps in each of the many otherwise required conventional HRTF filters. The 15 saved processing power can be used for other purposes, e.g., to enhance the HRTF effects.

Since ITDs are extracted, processed, and implemented separately in a roughly resolved delay module (i.e., the integer delay module 250), and in a finely tuned delay module (i.e., the fractional delay 20 module 252), the 3D audio effects can be easily controlled and adjusted to suit other special requirements, e.g., to be optimized for different head sizes. The super resolution sub-sample filtering polyphase filter based delay lines in accordance with the principles of the present invention introduce necessary delay without introducing discontinuity or 'clicks' in 25 the presentation to the listener.

The principles of the present invention are applicable for use in any 3D audio system that uses an interaural time delay as a localization queue for perceived direction of the sound by the listener. For instance, the present invention relates to 3D sound positioning in gaming, 30 virtualizing multiple loudspeaker array systems having two physical

speakers in AC3/Dolby™ Digital systems, advanced computer user interfaces, virtual acoustic reality software for architectural walk-throughs, auralization hardware/software, 3D enhancement for general stereo and wireless headphone sets, etc.

5 While the invention has been described with reference to the exemplary embodiments thereof, those skilled in the art will be able to make various modifications to the described embodiments of the invention without departing from the true spirit and scope of the invention.

## CLAIMS

What is claimed is:

1. A digital delay line for use in a 3D audio sound system,  
5 comprising:
  - a first delay module providing a choice of any delay within a first resolution; and
  - a second delay module in series with said first delay module, said second delay module providing a choice of any of a plurality of  
10 additional fractional delays, each of said additional fractional delays being less than said first resolution.
2. The digital delay line for use in a 3D audio sound system according to claim 1, wherein said first delay module comprises:  
15 a first-in, first out buffer.
3. The digital delay line for use in a 3D audio sound system according to claim 1, wherein said second delay module comprises:
  - a choice of any one of a plurality of polyphase filters, each of  
20 said polyphase filters providing an additional fraction delay less than said first resolution.
4. The digital delay line for use in a 3D audio sound system according to claim 1, further comprising:  
25 a localization control module comprising an interaural time delay look-up table associating desired sound source locations with a particular interaural time delay.

5. The digital delay line for use in a 3D audio sound system according to claim 4, wherein said localization control module further comprises:

an integer and fractional delay selector adapted to  
5 determine a first time delay for use by said first delay module and said additional fractional delay for use by said second delay module.

6. The digital delay line for use in a 3D audio sound system according to claim 1, wherein:

10 said first resolution is based on a sampling rate of a digital audio signal.

7. A method for providing an interaural time delay in a digital 3D sound system, comprising:

15 selecting one of a plurality of available first time delays having a first resolution between each of said plurality of available first time delays;

20 additionally selecting one of a plurality of available second time delays, each of said plurality of available second time delays being less than said first resolution; and

adding said selected first time delay and said second time delay to provide a desired interaural time delay.

8. The method for providing an interaural time delay in a  
25 digital 3D sound system according to claim 7, wherein:

said desired interaural time delay relates to a desired interaural time delay for one ear of a listener; and

said first time delay relates to a desired interaural time delay for a second ear of said listener.

9. The method for providing an interaural time delay in a digital 3D sound system according to claim 7, wherein:

    said plurality of available time delays are based on a sampling rate of a digital audio signal.

5

10. The method for providing an interaural time delay in a digital 3D sound system according to claim 7, further comprising:

    fixing a first interaural time delay with respect to a first ear of a listener; and

10           providing said desired interaural time delay with respect to a second ear of said listener.

11. Apparatus for providing an interaural time delay in a digital 3D sound system, comprising:

15           means for selecting one of a plurality of available first time delays having a first resolution between each of said plurality of available first time delays;

    means for additionally selecting one of a plurality of available second time delays, each of said plurality of available second 20 time delays being less than said first resolution; and

    means for adding said selected first time delay and said second time delay to provide a desired interaural time delay.

12. The apparatus for providing an interaural time delay in a 25 digital 3D sound system according to claim 11, wherein:

    said desired interaural time delay relates to a desired interaural time delay for one ear of a listener; and

    said first time delay relates to a desired interaural time delay for a second ear of said listener.

30

13. The apparatus for providing an interaural time delay in a digital 3D sound system according to claim 11, wherein:

    said plurality of available time delays are based on a sampling rate of a digital audio signal.

5

14. The apparatus for providing an interaural time delay in a digital 3D sound system according to claim 11, further comprising:

    means for fixing a first interaural time delay with respect to a first ear of a listener; and

10       means for providing said desired interaural time delay with respect to a second ear of said listener.

15

## ABSTRACT

A digital 3D sound audio source is implemented for digital audio using interaural time delays formed from two delay lines: a first delay line providing a rough estimate of the desired interaural time delay 5 for a particular audio sample, and a second delay line in series with the first delay line providing a more finely resolved delay. The use of the second delay line eliminates the need for conventional real-time interpolation techniques to provide the appropriate interaural time delay. In the disclosed embodiment, the first delay module, i.e., the integer delay 10 module, is formed from a first-in, first-out (FIFO) buffer with appropriate selection control of a desired sample as it passes through the FIFO buffer with each clock cycle based on the sampling rate. The second delay module (i.e., the fractional delay module) is formed from a plurality of polyphase (FIR) filters. The number of polyphase filters is determined 15 based on the desired resolution of the interaural time delay.

20

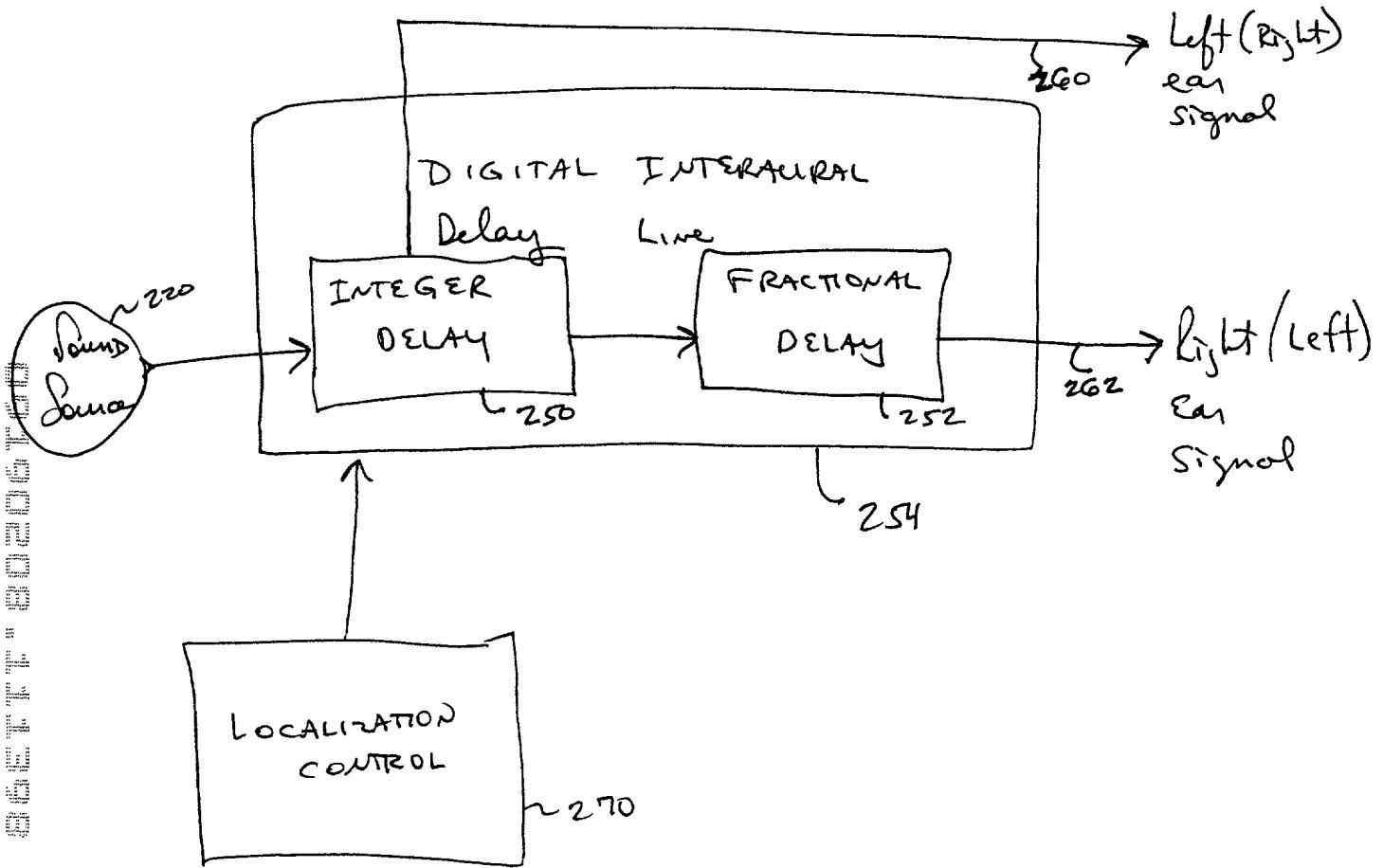


FIG. 1

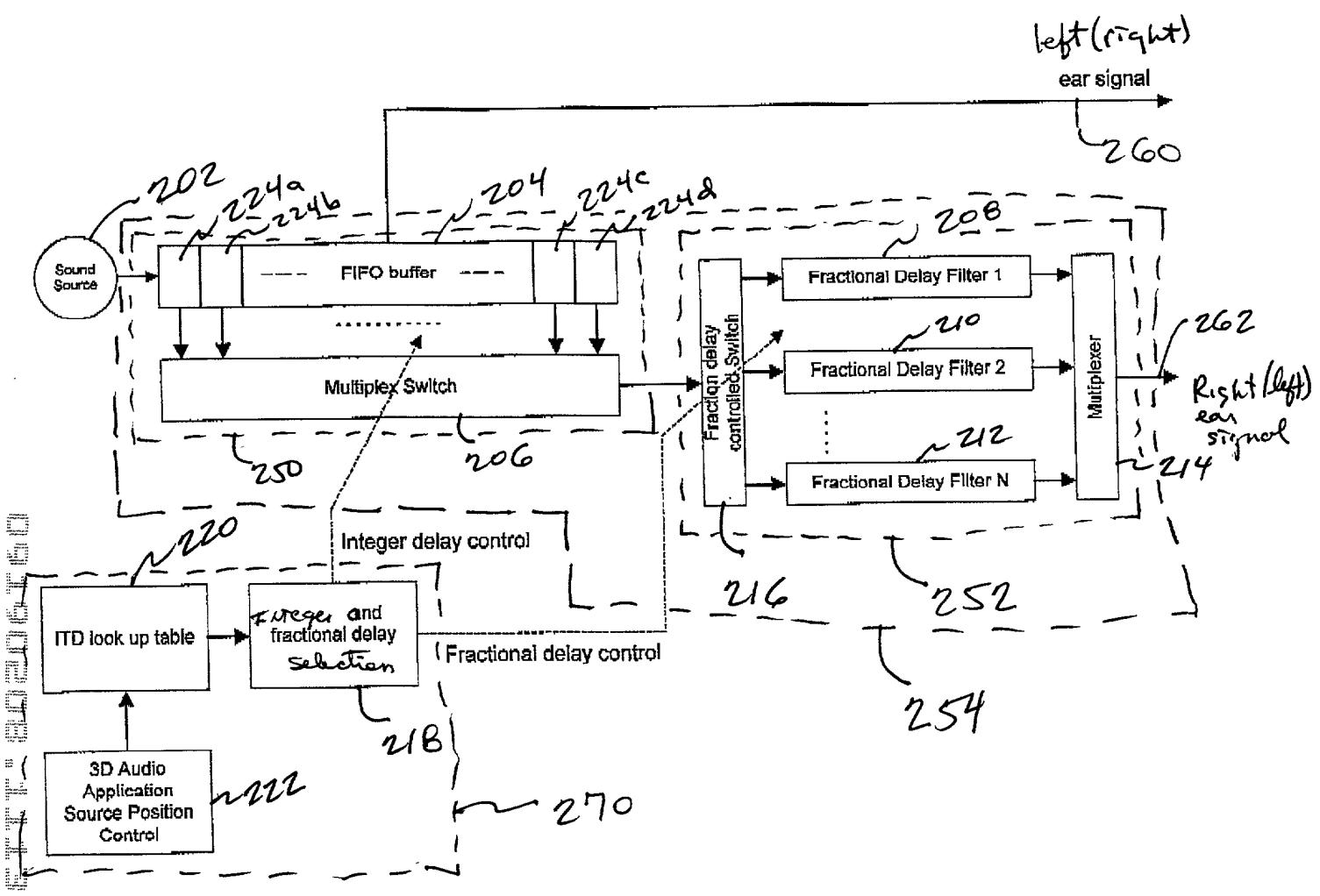
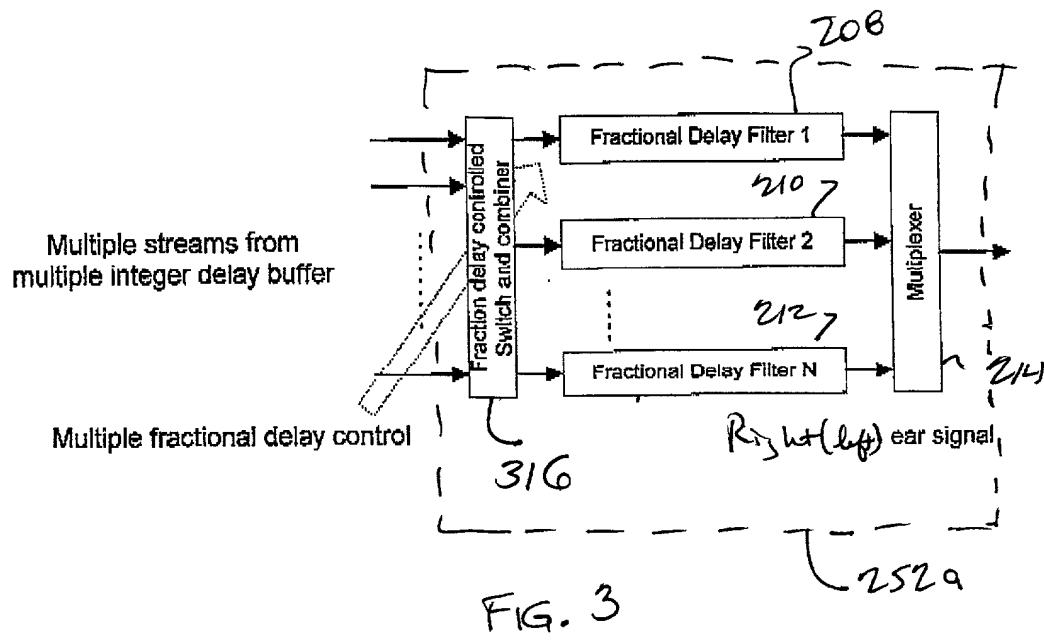


FIG. 2



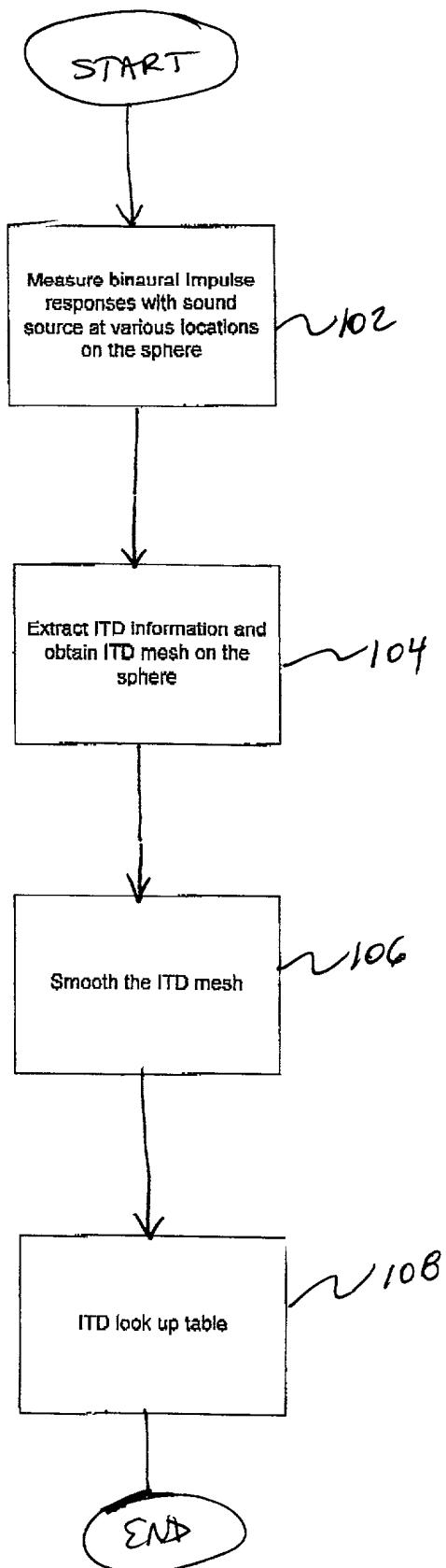


FIG. 4

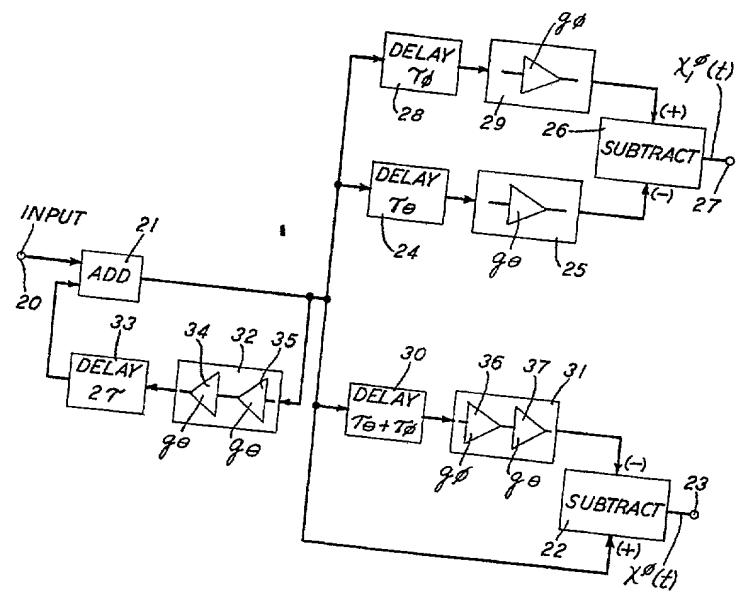


FIG. 5

Prior Art

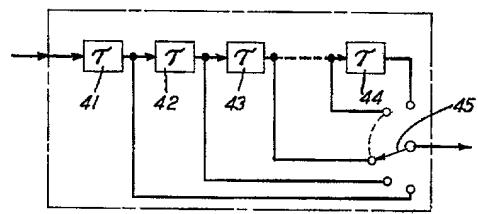
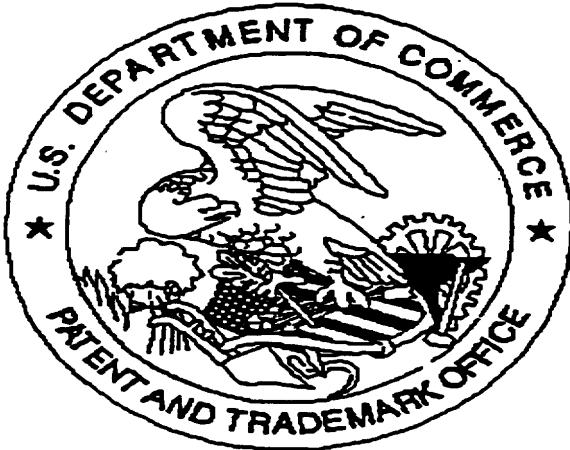


FIG. 6

PRIOR ART

United States Patent & Trademark Office  
Office of Initial Patent Examination – Scanning Division



Application deficiencies were found during scanning:

Page(s) \_\_\_\_\_ of \_\_\_\_\_ were not present  
for scanning. (Document title)

Page(s) \_\_\_\_\_ of \_\_\_\_\_ were not present  
for scanning. (Document title)

Scanned copy is best available.

SCANNED